RESEARCH

UDOT Keeping Utah Moving

Utah Department of Transportation - Research Division
4501 South 2700 West - P.O. Box 148410 - SLC,UT 84114-8410

Report No. UT-19.22

USING ATSPM DATA FOR TRAFFIC DATA ANALYTICS

Prepared For:

Utah Department of Transportation
Research & Innovation Division

Submitted By:

Resource System Group

Authored By:

Robert Chamberlin
Kiavash Fayyaz

Final Report
September 2019

DISCLAIMER

The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind, and assumes no liability therefore.

ACKNOWLEDGMENTS

The authors acknowledge the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Nicolas Black
- Mark Taylor
- Kevin Nichol
- Shana Lindsey

TECHNICAL REPORT ABSTRACT

1. Report No. UT-19.22	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A
4. Title and Subtitle Using ATSPM Data for Traffic Data Analytics		5. Report Date September 2019
		6. Performing Organization Code
7. Author(s) Robert Chamberlin, Kiavash Fayyaz		8. Performing Organization Report No.
9. Performing Organization Name and Address Resource Systems Group Crane Building, 307 W 200 S, Suite 2004 Salt Lake City, Utah 84101		10. Work Unit No. 5H08265H
		11. Contract or Grant No. 19-8273
12. Sponsoring Agency Name and Address Utah Department of Transportation 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410		13. Type of Report & Period Covered Final August 2018 to September 2019
		14. Sponsoring Agency Code PIC No. UT18.702

15. Supplementary Notes

Prepared in cooperation with the Utah Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration

16. Abstract

UDOT performs about 6000 short-duration counts over a three-year cycle to estimate and report AADT on roadways per FHWA requirements. Many of these roadways are in proximity to signalized intersections equipped with radar detectors that provide approach and turning movement counts. This research investigates the use of volume data obtained from traffic-signal radar detectors (i.e. Advance and Matrix detectors) to estimate AADTs and related traffic engineering factors. An assessment of the accuracy of these radar detectors may enable the elimination of selected short-duration counts, and possibly complement CCS data for estimating seasonal factors. In this research, 27 Matrix detectors and 33 Advance detectors proximate to CCS sites were identified. The hourly count data for an entire year, 2017, was collected from ATSPM data archive and mapped with the associated CCS hourly counts as ground-truth. An anomaly detection method was implemented to clean the dataset of count data when significant outliers were identified. The accuracy of detector hourly counts was measured using linear regression with and without adjustment factors. The results show that hourly counts from Matrix detectors hourly are more accurate (i.e. average R-squared value of 0.93) than Advance detectors' hourly counts (i.e. average R-squared value of 0.79). AADTs estimated from Matrix detectors had an 88 percent accuracy, with a range of -21%-+7%. Matrix detectors are sufficiently accurate for estimating AADT as the current methods utilizing short-duration counts have been estimated to be less than 80% accurate. The Matrix detectors are also very accurate in estimating the seasonal factors (i.e. about 97% accurate) and thus can be used to complement CCSs in calculating them. This would be particularly valuable to UDOT in measuring seasonal factors for lower functional class roadways which have sparse coverage by CCS sites.

17. Key Words Select specific and precise terms or phrases that identify principal subjects covered in the report and separate these by commas.		18. Distribution Statement Not restricted. Available through: UDOT Research Division 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410 www.udot.utah.gov/go/research		23. Registrant's Seal N/A
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 53	22. Price N/A	

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

UNIT CONVERSION FACTORS

SI* (MODERN METRIC) CONVERSION FACTORS				
APPROXIMATE CONVERSIONS TO SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

LIST OF ACRONYMS

AADT	Annual Average Daily Traffic
ATSPM	Automated Traffic-Signal Performance Measures
CCS	Continuous Count Station
FHWA	Federal Highway Administration
IQR	Inter Quantile Range
MADT	Monthly Average Daily Traffic
TAC	Technical Advisory Committee
TOD	Time of Day
UDOT	Utah Department of Transportation

EXECUTIVE SUMMARY

UDOT performs 6000 short-duration traffic counts over a three-year period to estimate and report AADT on roadways per FHWA requirements. Short duration counts range from 1-7 days, with 48-hour durations being typical. Many of these roadways are in proximity to signalized intersections equipped with radar detectors that estimate approach volumes.

This research investigates the use of data from traffic signal radar detectors (i.e. Wavetronix Advance and Matrix detectors) to estimate AADT, possibly eliminating the need for some short-duration counts, and for use in complementing CCS data for estimating seasonal adjustment factors. CCS traffic count data are obtained from two detector types: 1) from Wavetronix HD radar detectors, which are mounted roadside with a “sidefire” orientation; and, 2) from in-pavement inductance loops.

To this end, 27 Matrix detectors and 33 Advance detectors in proximity to CCS sites were identified. The hourly count data for the entire 2017 year was collected from ATSPM archive and mapped to the associated CCS hourly counts as ground-truth.

The hourly count data from the two detector types were “cleaned” to account for anomalous data. Three anomaly detection methods were implemented and their performance was compared. The TOD & IQR anomaly detection method showed superior performance in identifying anomalous hourly counts. Anomaly detection must be applied to both Matrix and Advance detectors’ hourly counts prior to any analysis since even a small number of large anomalies can have a significant impact on the accuracy of results.

Two linear regressions relating CCS hourly counts to detector hourly counts, one with no adjustment factor and a second with an adjustment factor, were implemented and R-squared and adjustment factor values were reported. Advance detectors’ hourly counts without adjustment factors have an average R-squared value of 0.79 but can substantially increase accuracy through the application of an adjustment factor (i.e. average R-squared value of 0.99). The Advance detectors are more accurate on roadways with two lanes compared to other lane configurations.

Matrix detector hourly counts are more accurate (i.e. average R-squared value of 0.93) than Advance detectors' hourly counts.

AADTs were calculated using the hourly counts from Matrix detectors and CCSs. The Matrix detectors estimated AADT with 88 percent accuracy compared to the AADTs estimated from CCS sites.

The Matrix detectors are also very accurate in estimating seasonal factors and thus can be used to complement CCS in calculating them. Matrix detectors estimate monthly seasonal factors with 97.5 percent accuracy and day of week in month seasonal factors with 96.8 percent accuracy for various functional classification groupings.

1.0 INTRODUCTION

1.1 Problem Statement

The Utah Department of Transportation (UDOT) is required to estimate and report AADT to the Federal Highway Administration (FHWA) every year on all federal-aid highways. For this purpose, UDOT has a significant investment in collecting short-duration counts (typically lasting two days) for approximately 6000 locations over each 3-year period (approximately 2000 locations per year). These counts are then used to estimate AADTs for the locations. There are several challenges in conducting short-duration counts as follows:

1. Costly and labor-intensive procedure,
2. Scheduling 2000 short-duration counts per year is challenging,
3. Safety risks associated with installing tube counters on high-speed road segments.

1.2 Objectives

UDOT's Automated Traffic Signal Performance Measures (ATSPM) include several hundred signalized intersections instrumented with radar-detection technology that, among other operational parameters, measures turning movement volumes for all or some intersection approaches. ATSPM is a rich data set with great potential to be used in broader traffic analytic applications. This research investigates the use of ATSPM radar detector data to estimate AADT (possibly eliminating the need for some short-duration counts) and to complement Continuous Counter Station (CCS) data in estimating the adjustment factors.

In describing the objectives of this research, it is important to cite two recent UTRAC studies, Saito, et al 2015 and Saito, et al 2016, "Calibration of Automatic Performance Measures – Speed and Volume Data, Volumes 1 and 2". The present research differs from the previous research in two important ways:

1. The Saito research used manual counts as the ground truth. The current study uses machine-detector counts obtained from Matrix and Advance detectors that have been matched with CCS sites. As a result, the "ground truth" in the present research is from another device and not a manual count. The present research is relevant in that FHWA accepts the machine counts from CCS sites (obtained from HD sensors or from in-pavement loops) as valid representations of traffic volume. This research seeks to quantify how accurate counts obtained from another device, namely, the Matrix and Advance detectors installed at several ATSPM sites, are when compared to the volumes from CCS sites already considered to be valid by FHWA.

2. The Saito research conducted ground truth counts manually at selected intersections for one-hour periods, then compared to corresponding volume counts obtained from the Matrix (Vol 1) and Advance (Vol 2) detectors. The present research analyzes volume data from ATSPM and CCS sites over an entire year, 2017 (8760 hours of data for the intersections that were generating counts over the entire year). This is a critical difference. Over the course of a year, factors may act to compensate due to the much larger data set.

1.3 Scope

TASK 0: Project Management and Steering Committee

A Steering Committee was assembled to guide the research and review deliverables. RSG held a kick-off meeting with the Technical Advisory Committee (TAC) on July 2, 2018. The main objective of the meeting was to specify deliverables for each specific task. Additional TAC meetings were held on February 11, 2019 and on May 6, 2019 to discuss the data collection, methodology and preliminary findings.

In addition to ongoing communication and TAC meetings, the research team provided periodic progress reports to the UDOT project champion, Nicolas Black, throughout the duration of the project.

TASK 1: Selection of Test Locations

Test ATSPM intersections were selected based on their proximity to existing UDOT CCS (as the ground-truth dataset). The procedure involved mapping CCS and ATSPM Matrix and Advance detectors' location on the same roadway segments with no inlet or outlet streets in between.

TASK 2: Prepare Data Sets

Task 2 focused on preparing the comparative data sets from CCS and ATPSM data for the intersections selected in Task 1. The ATPSM and CCS hourly count data were collected for each location for the entire 2017 year.

TASK 3: Analysis of ATSPM Data Compared with CCS

Task 3 focused on comparing the count datasets for the test ATSPM locations that are associated with UDOT CCSs. Our statistical analysis focused on five items:

1. Map the collected ATSPM and CCS data

2. Detect anomalies (from device malfunction or other) in ATSPM data

3. Determine the accuracy of ATSPM hourly counts vs. CCS hourly counts

4. Determine the accuracy of Matrix and Advance detector AADT vs. CCS's AADT

5. Determine the accuracy of Matrix detectors in estimating seasonal factors

TASK 3: Final Report, Findings & Recommendations

In this task the results of the data collection and analysis are summarized, and overall accuracy was calculated for each ATSPM detection device. The resulting research report includes results of the site selection, data matching, and detection accuracy.

1.4 Outline of Report

This report documents the findings of the research and includes the following chapters:

- Introduction
- Site Selection
- Prepare Data Set
- Analysis of ATSPM Data Compared with CCS
- Conclusions
- Recommendations and Implementation

2.0 SITE SELECTION

2.1 Overview

This chapter describes the procedure on selecting test locations. The test sites were selected by finding the ATSPM detectors in proximity to CCSs. We first describe the two types of ATSPM detectors considered in this study, and then provide their matching procedure to CCSs.

2.2 ATSPM Detector Types

There are two types of ATSPM detectors analyzed in this study as follows:

1. Wavetronix SmartSensor Advance (referred to as Advance)

2. Wavetronix SmartSensor Matrix (referred to as Matrix)

Wavetronix SmartSensor Advance detectors are located approximately 300-400 ft upstream of signalized intersections and provide total through traffic counts. Wavetronix SmartSensor Matrix detectors are located at signal stop-bar locations and provide lane-by-lane turning movement counts. For information about these detectors' technology, please refer to chapter 2 of UTRAC project no. UT-16.05 (Saito et al. 2016).

2.3 ATSPM Detector Locations and Matching Process with CCS

UDOT provided two ArcMap shapefiles including a shapefile of all signals equipped with Matrix detectors (almost all signals equipped with Matrix detectors are also equipped with Advance detectors) and a shapefile of all CCSs. The test sites were selected as locations with paired CCS and Matrix detectors (or Advance detectors) that were 1) in close proximity of each other with no or limited traffic access points in between, 2) estimated volume counts for the same route and direction, and 3) having the count data available for the same period.

As of August 2018, there were 962 signals equipped with Matrix detectors and 113 CCSs in the state of Utah. The complete list of CCS and Matrix detector pairs that provide the counts for the same route, direction, and time (i.e. entire 2017) are provided in Table 2.1, showing a total of eleven CCS sites that are successfully paired with ATSPM signals outfitted with 27 Matrix detectors.

Table 2.1 Matched CCS and Matrix Detector List

CCS Number	CCS Direction	Signal ID	Direction	Primary Name	Secondary Name
-316	Negative	5162	Southbound	US-89	Cherry Lane (Layton)
-316	Positive	5162	Northbound	US-89	Cherry Lane (Layton)
-333	Negative	7185	Southbound	700 East	1300 South
-333	Positive	7185	Northbound	700 East	1300 South
-335	Negative	7335	Westbound	5300 South	700 West
-335	Positive	7335	Eastbound	5300 South	700 West
-354	Negative	7291	Westbound	3300 South	900 West
-354	Positive	7291	Eastbound	3300 South	900 West
-354	Negative	7502	Westbound	3300 South	1200 West
-354	Positive	7502	Eastbound	3300 South	1200 West
-355	Negative	7275	Westbound	3500 S (SR-171)	8000 W
-355	Positive	7275	Eastbound	3500 S (SR-171)	8000 W
-402	Negative	8150	Westbound	SR-9 (State St)	6300 W (Telegraph Old Hwy 91)
-402	Positive	8150	Eastbound	SR-9 (State St)	6300 W (Telegraph Old Hwy 91)
-406	Negative	7193	Southbound	4500 South	700 East
-406	Positive	7193	Northbound	4500 South	700 East
-406	Negative	7211	Eastbound	Van Winkle (SR-152)	900 East
-406	Positive	7211	Westbound	Van Winkle (SR-152)	900 East
-407	Negative	7391	Southbound	SR-68 (Redwood Rd)	14400 South
-407	Positive	7391	Northbound	SR-68 (Redwood Rd)	14400 South
-620	Negative	5320	Westbound	1000 West	200 North (Logan)
-620	Positive	5320	Eastbound	1000 West	200 North (Logan)
-626	Negative	6145	Westbound	SR-73	Foothill Blvd.
-626	Positive	6145	Eastbound	SR-73	Foothill Blvd.
-626	Negative	6185	Westbound	Cory Wride Hwy (SR-73)	Mt Airey Dr
-631	Negative	7313	Westbound	4500 South	815 West
-631	Positive	7313	Eastbound	4500 South	815 West

The complete list of CCS and Advance detector pairs that provide the counts for the same route, direction, and time (i.e. entire 2017) are provided in Table 2.2. The list shows 23 CCS sites that are paired with 33 Advance detectors.

Table 2.2 Matched CCS and Advance Detectors List

CCS Number	CCS Direction	Signal ID	Direction	Primary Name	Secondary Name
-622	Positive	5297	Northbound	Main St. (SR-165)	1700 S (Providence)
-620	Negative	5320	Westbound	1000 West	200 North (Logan)
-329	Negative	5011	Southbound	Washington	Riverdale Rd
-316	Negative	5162	Southbound	US-89	Cherry Lane (Layton)
-332	Negative	7219	Northbound	Foothill Drive	Sunnyside
-333	Negative	7185	Southbound	700 East	1300 South
-408	Negative	7093	Southbound	Redwood Road	California (1300S)
-354	Negative	7502	Westbound	3300 South	1200 West
-355	Negative	7275	Westbound	3500 S (SR-171)	8000 W
-406	Positive	7193	Northbound	4500 South	700 East
-631	Positive	7313	Eastbound	4500 South	815 West
-335	Negative	7347	Westbound	5400 South	1070 West
-634	Negative	7524	Southbound	SR-85 SB (Mtn View)	7800 South
-407	Negative	7392	Southbound	Redwood Rd.	Porter Rockwell Blvd
-633	Positive	7507	Westbound	SR-85 NB (Mtn View)	Porter Rockwell Blvd
-509	Positive	6100	Southbound	US-40	SR-32 / River Road
-626	Positive	6145	Eastbound	SR-73	Foothill Blvd.
-632	Positive	6035	Eastbound	Pioneer Crossing	Millpond Drive
-319	Negative	6425	Southbound	University Avenue	SR-52 (800 N Orem)
-350	Positive	6420	Northbound	University Avenue	3300 North
-425	Negative	6256	Westbound	US-40	2000 W
-402	Positive	8151	Eastbound	SR-9 (State St)	5300 West (SR-318)
-332	Positive	7503	Southbound	Foothill Drive	2100 East
-354	Positive	7291	Eastbound	3300 South	900 West
-631	Negative	7312	Westbound	4700 South (SR-266)	Atherton Dr (1050 W)
-335	Positive	7335	Eastbound	5300 South	700 West
-633	Negative	7508	Southbound	SR-85 SB (Mtn View)	Porter Rockwell Blvd
-626	Negative	6185	Westbound	Cory Wride Hwy (SR-73)	Mt Airey Dr
-402	Negative	8150	Westbound	SR-9 (State St)	6300 W (Telegraph Old Hwy 91)
-408	Positive	7092	Northbound	Redwood Rd (SR-68)	Indiana Ave (850 S)
-406	Negative	7211	Eastbound	Van Winkle (SR-152)	900 East
-407	Positive	7391	Northbound	SR-68 (Redwood Rd)	14400 South
-308	Negative	6105	Northbound	Main St (Heber)	US-189 (1200 S)

The ATSPM system records each signal detector's data on an assigned "Detector Channel". Table 2.3 shows the detector channels assigned to Matrix and Advance detectors' records for signal ID 5162 as an example. Note that the Matrix detector provides lane-by-lane counts for this signal for all directions. For example, the Northbound through movement counts are recorded on detector channels 21 and 22 for the two Northbound through lanes. The Advance detector provides directional counts only for Northbound and Southbound (no Advance detectors for Westbound and Eastbound) and they are recorded on detector channels 2 and 4, respectively.

2.4 Summary

The test sites were selected as locations with pairs of CCS and Matrix detectors (or Advance detectors) that are 1) in close proximity of each other with no or limited inbound/outbound traffic in between, 2) providing the counts for the same route and direction, and 3) having the count data available for the same period. A total of 27 CCS-Matrix detector pairs (representing eleven discrete CCS sites) and 33 CCS-Advance detector pairs (representing 23 discrete CCS sites) that provide the counts for the same route, direction, and time were identified.

Table 2.3 Detector Data Labels for Signal ID 5162

Signal ID	Primary Name	Secondary Name	Detector Channel	Direction	Movement	Lane Number	Detector Type
5162	US-89	Cherry Lane (Layton)	2	Northbound	Thru	1	Advanced Count
5162	US-89	Cherry Lane (Layton)	4	Southbound	Thru	1	Advanced Count
5162	US-89	Cherry Lane (Layton)	20	Northbound	Left	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	21	Northbound	Thru	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	22	Northbound	Thru	2	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	23	Northbound	Right	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	31	Southbound	Left	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	32	Southbound	Thru	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	33	Southbound	Thru	2	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	34	Southbound	Right	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	41	Eastbound	Left	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	42	Eastbound	Thru-Right	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	48	Westbound	Left	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	49	Westbound	Thru	1	Lane-by-lane Count (Matrix)
5162	US-89	Cherry Lane (Layton)	50	Westbound	Right	1	Lane-by-lane Count (Matrix)

3.0 PREPARE DATA SETS

3.1 Overview

The data collection process included collecting traffic count data from the specified detector channels in 15-minute count bins and CCS hourly counts for the selected sites. The following sections describe the collected CCS and ATSPM count data.

3.2 CCS Data

The CCS count data was provided by Nicolas Black, UDOT Traffic Analysis Supervisor. The dataset includes the hourly directional counts for 113 CCSs for every day of 2017 which translates into 1,846,104 unique hourly directional counts. Figure 3.1 shows the 113 CCS locations in the state of Utah.

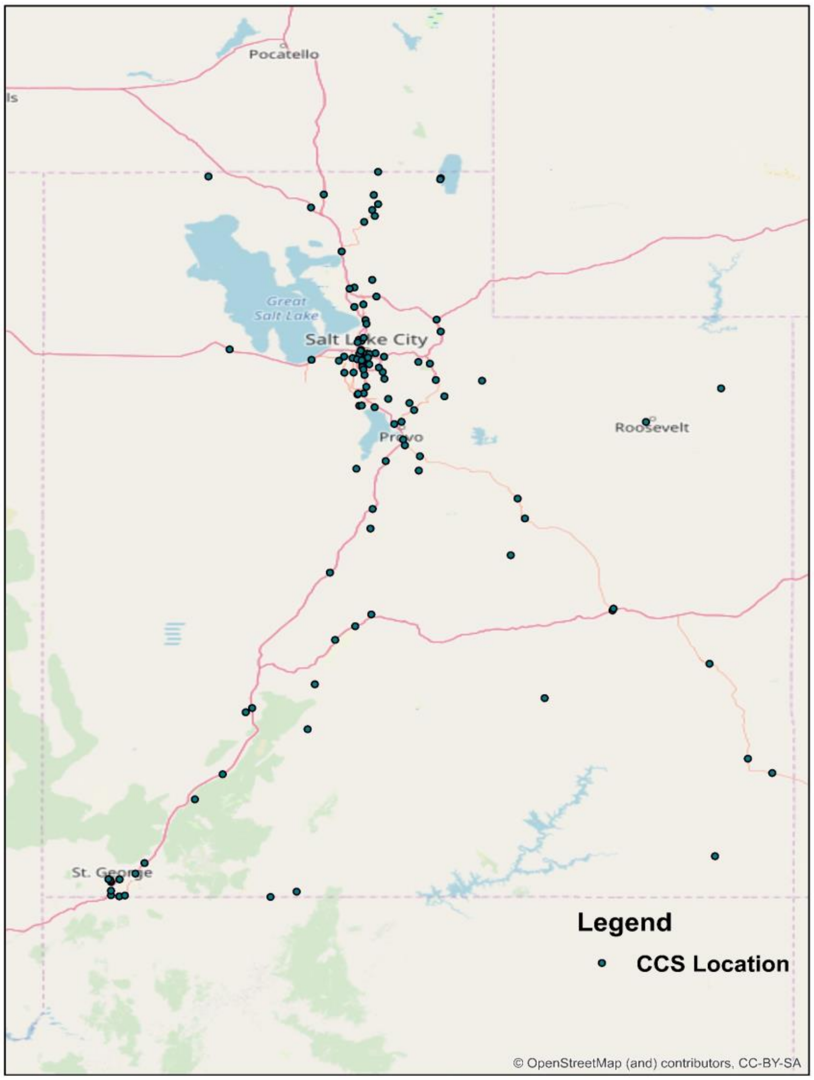

Figure 3.1 CCS Locations in the State of Utah

A sample of hourly directional CCS count data between 00:00 AM to 01:00 AM for the first two days of 2017 for CCS ID -0301 is shown in Table 3.1. The CCS ID -0301 is located one mile north of the I-215 and I-80 intersection at Parleys Canyon, on I-80.

Table 3.1 Sample Count Data for CCS ID -0301

CCS Number	Date	CCS Direction	Hour	Count
-0301	1/1/2017	Pos	0	373
-0301	1/1/2017	Neg	0	503
-0301	1/2/2017	Pos	0	136
-0301	1/2/2017	Neg	0	218

3.3 Matrix and Advance Detector Data

The ATSPM location and count data was provided by Mark Taylor and Jamie Mackey, of UDOT's Traffic Management Division. The dataset includes the 15-minute detector channel counts for 47 signals for every day of 2017 which translates into 15,999,450 unique 15-minute detector channel counts. Figure 3.2 shows the 965 signals with Matrix detectors in the state of Utah.

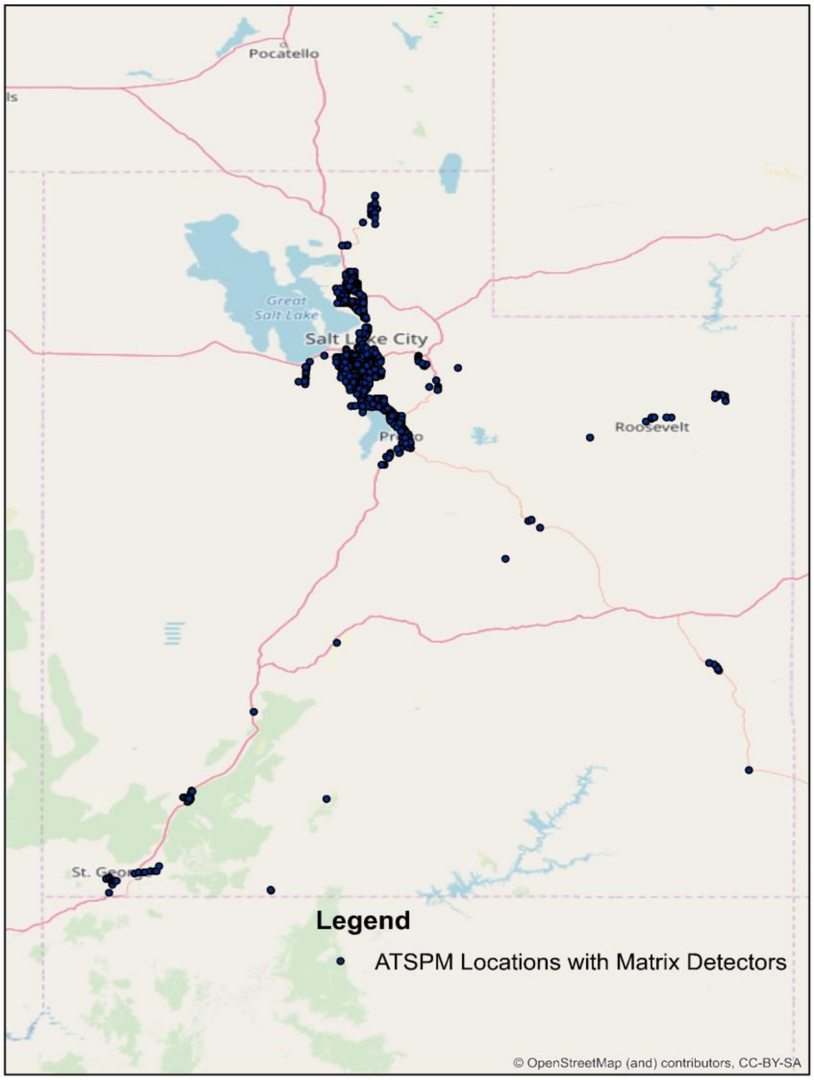

Figure 3.2 Signals with Matrix Detectors Locations in the State of Utah

Table 3.2 shows samples of detector channels 2, 4, 21, and 22 data for signal ID 5162 (US89 @ Cherry Lane, Layton). For example, the Matrix detector counts 22 + 33 (detector channels 21 and 22) vehicles moving in a Northbound through direction from 00:00 AM to 00:15 AM for signal ID 5162.

Table 3.2 Sample Detector Channels Number 2, 4, 21, and 22 Data for Signal ID 5162

Signal ID	Detector Channel	Time & Date	Volume
5162	2	2017-01-01T00:00:00Z	55

5162	4	2017-01-01T00:00:00Z	37
5162	21	2017-01-01T00:00:00Z	22
5162	22	2017-01-01T00:00:00Z	33
5162	2	2017-01-01T00:15:00Z	108
5162	4	2017-01-01T00:15:00Z	145
5162	21	2017-01-01T00:15:00Z	57
5162	22	2017-01-01T00:15:00Z	66
5162	2	2017-01-01T00:30:00Z	114
5162	4	2017-01-01T00:30:00Z	182
5162	21	2017-01-01T00:30:00Z	46
5162	22	2017-01-01T00:30:00Z	60
5162	2	2017-01-01T00:45:00Z	108
5162	4	2017-01-01T00:45:00Z	131
5162	21	2017-01-01T00:45:00Z	41
5162	22	2017-01-01T00:45:00Z	64
5162	2	2017-01-01T01:00:00Z	102
5162	4	2017-01-01T01:00:00Z	97
5162	21	2017-01-01T01:00:00Z	17
5162	22	2017-01-01T01:00:00Z	29

3.4 Summary

The CCS dataset includes the hourly directional counts for 113 CCSs for every day of 2017 which translates into 1,846,104 unique hourly directional counts. The ATSPM dataset, provided by UDOT's Traffic Management Division, includes the 15-minute detector channel counts for 47 ATSPM signals for every day of 2017 which translates into 15,999,450 unique 15-minute detector channel counts.

The traffic count data from the matched CCS-ATSPM pairs were assembled for identical time slices for statistical analysis of comparability in the following task.

4.0 Analysis of ATSPM Data Compared with CCS

4.1 Overview

This chapter focuses on the mapping and analyzing of Matrix and Advance detector data with the associated CCS data. A total of eleven CCS sites were successfully associated with

ATSPM signals with Matrix detectors and 23 CCS sites were associated with ATSPM signals with Advance detectors. This chapter describes an anomaly detection method utilized to filter the detector data, the accuracy of Matrix and Advance detectors when compared to the CCS count data, and how they can complement CCS data.

The analysis procedure involved the following five steps:

1. Map the collected ATSPM and CCS data;
2. Detect anomalies (caused by device malfunction or other) in ATSPM data and filter the data to eliminate outliers;
3. Determine the accuracy of ATSPM hourly counts vs. CCS hourly counts;
4. Determine the accuracy of AADTs estimated from Matrix and Advance detectors vs. AADTs estimated at CCS sites;
5. Determine the accuracy of Matrix detectors in estimating seasonal adjustment factors.

4.2 Map the Collected ATSPM and CCS Data

The Matrix and Advance detector data must first be mapped to the paired CCS data. The first step in the data mapping process is to use unified direction labeling. The CCS data labels directions as positive or negative according to increasing or decreasing milepost direction, whereas ATSPM data labels direction as Northbound, Southbound, Eastbound, or Westbound. To achieve unified labeling, the CCS directions were converted to Northbound, Southbound, Eastbound, or Westbound labels.

The second step in the data mapping process is to use a unified time interval (aggregation period) for all data sets. The CCS counts are provided in one-hour intervals, whereas the ATSPM counts are in 15-minute intervals. To achieve unified time interval, each four 15-minute ATSPM counts in an hour were summed to provide the counts in hour intervals.

The third step in the data mapping process is to find detector channels (lanes) that represent the entire CCS directional counts. Figure 4.1 shows the detectors and CCS positions and mapping for Foothill Drive and Sunnyside Avenue intersection. The northbound Advance detector is installed 370 ft. upstream (to the south) of the intersection and records all northbound counts in one detector channel. Thus, it is easy to match northbound CCS counts with Advance detector counts. However, the southbound Advance detector is installed 410 ft. upstream (to the north of the intersection) and thus cannot be matched to the southbound direction of CCS because of the intersection in between.

The Matrix detectors are installed at the intersection and provide lane-by-lane counts. In the case of the Foothill Drive and Sunnyside Avenue intersection, the northbound CCS counts must be matched with the sum of three northbound through lanes, one northbound left lane, and one northbound right lane. The southbound CCS counts must be matched with the sum of three southbound through lanes, one eastbound right lane, and two westbound left lanes. Note that this directional match is possible due to the availability of Matrix "lane-by-lane counts" for all directions. In other words, if the westbound and eastbound directions do not have Matrix detectors at this intersection, it would not be possible to accurately match the southbound CCS traffic flow for this location.

Of the eleven paired CCS-Matrix sites selected for this research, seven have cases where Matrix detectors enumerate approach counts from through-right lanes. The combination of movements will inevitably create an inaccuracy when adding the total volume from the detector count to the "exit volume" for that intersection.

Figure 4.1 Directional Movement Mapping for Matrix Detector, Advance Detector, and CCS at Foothill Dr. and Sunnyside Ave. Intersection

Finally, the Matrix and Advance detectors' counts are mapped to paired CCS counts according to matching CCS Number-ATSPM Signal ID, CCS Direction-Matrix/Advance Detector Channel, date, and time. Table 4.1 shows a sample of mapped count data collected from a Matrix detector, an Advance detector, and a CCS for ATSPM signal ID 5162 in southbound direction.

Table 4.1 Sample Mapped Matrix, Advance, and CCS Counts

Signal ID	Direction	CCS Number	CCS Direction	Date	Hour	Matrix Count	Advance Count	CCS Count
5162	Southbound	-316	Negative	1/1/2017	0	429	495	512
5162	Southbound	-316	Negative	1/1/2017	1	220	250	237
5162	Southbound	-316	Negative	1/1/2017	2	97	101	91
5162	Southbound	-316	Negative	1/1/2017	3	58	70	67
5162	Southbound	-316	Negative	1/1/2017	4	60	71	73
5162	Southbound	-316	Negative	1/1/2017	5	77	111	104
5162	Southbound	-316	Negative	1/1/2017	6	113	148	146
5162	Southbound	-316	Negative	1/1/2017	7	131	165	163
5162	Southbound	-316	Negative	1/1/2017	8	257	287	291
5162	Southbound	-316	Negative	1/1/2017	9	268	325	321

4.3 Detect and Remove Anomalies

Once the count data were mapped, the Matrix and Advance detector data were cleaned to remove anomalous data. These anomalies may be caused by device malfunction, change in device orientation due to extreme weather, or other unknown reasons. Figure 4.2 shows the southbound direction Advance detector counts and CCS counts for signal ID 6100 for all of 2017. It appears that the Advance detector is malfunctioning in May, November, and December. These anomalies must be identified and removed to achieve quality detector count data.

Three different anomaly detection methods were tested that are described in the following subsections. The key characteristic of the anomaly detection methods is that they should be solely based on detectors' data. This ensures that the method is applicable to other places where no ground-truth (i.e. no CCS) data is available.

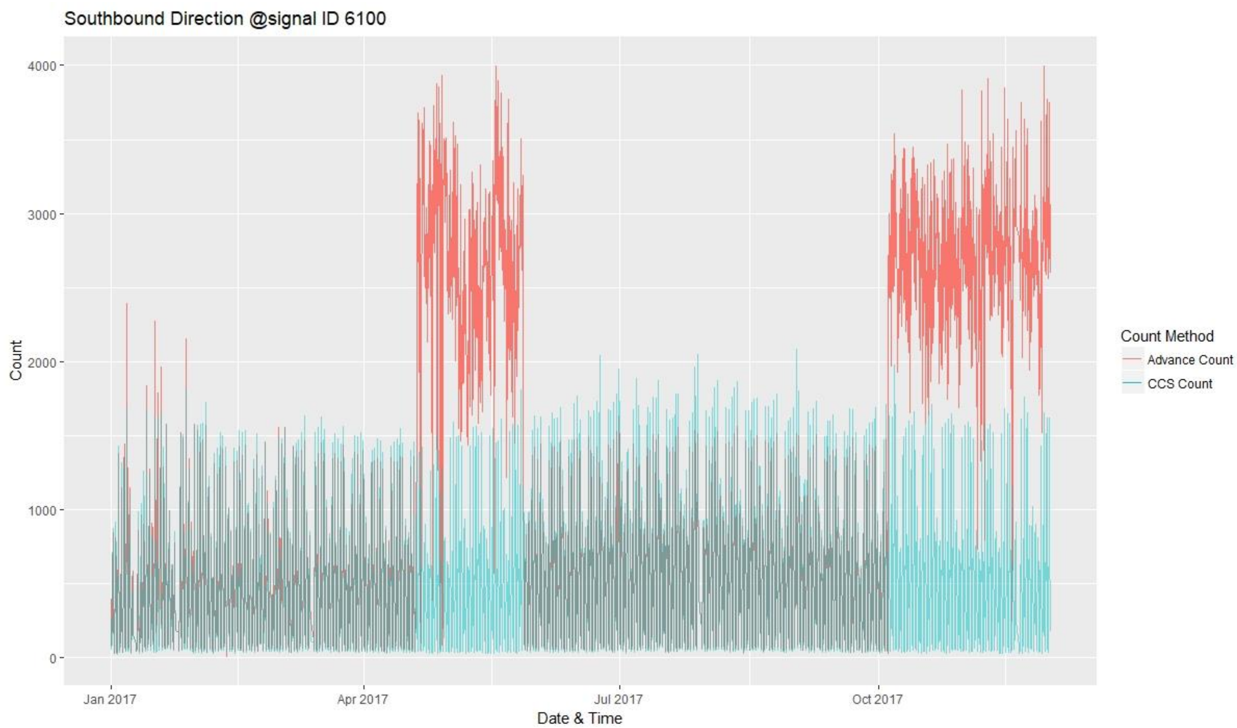

Figure 4.2 Southbound Direction Advance Detector Counts and CCS Counts for Signal ID 6100 for Entire 2017

4.3.1 Inter Quantile Range-Based Anomaly Detection

The Inter Quantile Range (IQR)-based anomaly detection method assumes that any hourly count data that is bigger than the 75th percentile of all hourly counts in the same hour + 1.5 * IQR is an outlier. The IQR is calculated as:

$$IQR_{hid} = 75th\ percentile_{hid} - 25th\ percentile_{hid} \qquad (4\text{-}1)$$

where IQR_{hid} is the IQR for hour *h* signal *i* on direction *d*, $75th\ percentile_{hid}$ is the 75th percentile of all hourly counts for hour *h* signal *i* on direction *d*, and $25th\ percentile_{hid}$ is the 25th percentile of all hourly counts for hour *h* signal *i* on direction *d*. The threshold is then calculated as follows:

$$Threshold_{hid} = 75th\ percentile_{hid} + 1.5 * IQR_{hid} \tag{4-2}$$

where $Threshold_{hid}$ is the anomaly detection threshold for hour *h* signal *i* on direction *d*, $75th\ percentile_{hid}$ is the 75th percentile of all hourly counts for hour *h* signal *i* on direction *d*, and where IQR_{hid} is the IQR for hour *h* signal *i* on direction *d*. Any hourly count value greater than or less than this threshold will be identified as an anomaly. The datasets used for this analysis were determined to have only high-end anomalies; no undercounting anomalies were identified in the dataset

4.3.2 K-Means Clustering Anomaly Detection

K-means clustering partitions hourly count observations into a desirable number of clusters in a way that each observation belongs to the cluster with the nearest average count value. For anomaly detection purposes, we will have two clusters, one with the anomalous observations and one with valid observations. Each hourly count observation will belong to the cluster that has the closest average hourly count value. For example, assume we have 200 hourly counts for signal *i* in direction *d* where 180 observations are in the valid cluster and 20 observations are in the anomaly cluster. Let's also assume that the average hourly counts are 300 and 700 for valid and anomaly clusters, respectively. Let's also assume that a new hourly count is collected (observation 201) with the count value of 340. The new observation will go to the valid category since its count is closer to the valid cluster average count than the anomaly cluster average count (i.e. $|340 - 300| < |340 - 700|$). Consequently, the valid cluster's average count will get updated (i.e. $\frac{200*300+340}{201} = 300.199$) since a new hourly count is added.

4.3.3 Time of Day and IQR-Based Anomaly Detection

Time of Day (TOD) and IQR-based method detects anomalies as:

1. All hourly counts in each day are anomalies if the count data from 01:00 AM to 03:00 AM are suspiciously high for that day,

2. An hourly count in the remining days is anomaly if the hourly count is greater than the threshold explained in IQR-based anomaly detection.

The threshold for the hourly counts from 01:00 AM to 03:00 AM are calculated as follows:

$$\begin{aligned} &if\ med_{123_{id}} > mean_{123_{id}} \rightarrow\ Threshold_{123_{id}} = med_{123_{id}} * 2 \\ &if\ med_{123_{id}} < mean_{123_{id}} \rightarrow\ Threshold_{123_{id}} = mean_{123_{id}} * 4 \end{aligned} \tag{4-3}$$

where $med_{123_{id}}$ and $mean_{123_{id}}$ are the median and mean of historic hourly counts at 1, 2, and 3 AM for signal *i* on direction *d*, respectively. $Threshold_{123_{id}}$ is the threshold for hourly counts at 1, 2, and 3 AM for signal *i* on direction *d*. All hourly counts in each day will be marked as anomaly if at least one of the hourly counts at 1, 2, and 3 AM is higher than this threshold.

4.3.4 Anomaly Detection Results

The results of anomaly detection algorithms were compared by calculating the accuracy (R-squared) of filtered datasets in measuring the actual counts. Table 4.2 shows the accuracy of IQR, K-means, and TOD & IQR anomaly detection methods on Advance detectors. Note that all the R-Squared values in Table 4.2 and Table 4.3 are reported based on adjustment factors derived from a comparison of the detector data against the CCS data for like movements. The adjustment factoring is described in more detail in Section 4.4. The TOD & IQR methods show the best performance by providing higher accuracy for almost all Advance detectors compared to unfiltered data, IQR method, and K-means method. It is important to note that anomaly detection has a significant impact on the accuracy of the results and thus must be implemented for Advance detectors.

Table 4.2 Accuracy of Three Anomaly Detection Methods and Percentage of Anomaly Observation in TOD & IQR Method for Advance Detectors

Signal ID	Direction	CCS Number	Unfiltered R2	Filtered R2			% Anomaly
				IQR	K-means	TOD & IQR	
5297	Northbound	-622	0.88	1.00	0.99	1.00	5.8%
5320	Westbound	-620	0.45	0.99	0.98	0.99	7.5%
5011	Southbound	-329	1.00	1.00	0.98	1.00	5.2%
5162	Southbound	-316	0.99	0.99	0.99	0.99	3.4%
7219	Northbound	-332	0.54	0.93	0.98	0.98	15.1%
7185	Southbound	-333	0.99	0.99	0.95	0.99	2.1%
7093	Southbound	-408	0.89	0.95	0.95	0.95	9.0%
7502	Westbound	-354	0.72	1.00	0.95	1.00	7.6%
7275	Westbound	-355	0.45	0.98	0.89	0.99	23.6%
7193	Northbound	-406	0.85	0.99	0.98	0.99	3.2%
7313	Eastbound	-631	1.00	1.00	0.99	1.00	3.8%
7347	Westbound	-335	0.86	0.97	0.94	0.98	13.9%
7524	Southbound	-634	0.92	0.99	0.98	0.99	2.6%
7392	Southbound	-407	0.98	0.98	0.98	0.98	2.3%
7507	Westbound	-633	0.35	0.35	0.90	0.97	48.1%
6100	Southbound	-509	0.47	0.47	0.97	0.99	30.8%
6145	Eastbound	-626	0.97	0.99	0.98	0.99	3.7%
6035	Eastbound	-632	0.23	0.99	0.87	0.99	5.2%
6425	Southbound	-319	0.29	0.59	0.59	0.89	41.5%
6420	Northbound	-350	0.99	0.99	0.95	0.99	4.9%
6256	Westbound	-425	0.99	0.99	0.98	0.99	1.6%
8151	Eastbound	-402	0.43	0.99	0.83	0.99	12.8%
7503	Southbound	-332	0.93	0.95	0.86	0.96	15.0%
7291	Eastbound	-354	0.55	0.88	0.79	0.90	11.1%
7312	Westbound	-631	1.00	1.00	1.00	1.00	2.1%
7335	Eastbound	-335	1.00	1.00	1.00	1.00	3.5%
7508	Southbound	-633	0.32	0.32	0.97	0.99	37.4%
6185	Westbound	-626	0.96	0.98	0.74	0.99	4.9%
8150	Westbound	-402	0.99	0.99	0.97	0.99	4.9%
7092	Northbound	-408	1.00	1.00	0.99	1.00	2.8%
7211	Eastbound	-406	0.98	0.98	0.97	0.98	3.0%
7391	Northbound	-407	0.99	0.99	0.99	0.99	2.2%
6105	Northbound	-308	0.84	0.86	0.87	0.86	4.2%
	Average Value:		**0.78**	**0.91**	**0.93**	**0.98**	**10.5%**

TOD & IQR anomaly-detection-method accuracy was validated by applying to Matrix detectors as well. Table 4.3 shows the accuracy of TOD & IQR anomaly detection method for

Matrix detectors. The results show that Matrix detectors (with average anomalous observation of 3.3 percent) are much less prone to malfunctioning compared to Advance detectors (with average anomalous observations of 10.5 percent).

Table 4.3 Accuracy of TOD & IQR Detection Method for Matrix Detectors

Signal ID	Direction	CCS Number	Unfiltered R2	Filtered R2 for TOD & IQR	% Anomaly
7313	Westbound	-631	1.00	1.00	1.5%
7313	Eastbound	-631	1.00	1.00	2.1%
6145	Westbound	-626	0.98	0.98	2.6%
6185	Westbound	-626	0.99	0.99	4.1%
6145	Eastbound	-626	0.99	0.99	3.0%
5320	Westbound	-620	0.98	0.98	2.4%
5320	Eastbound	-620	0.99	0.99	4.6%
7391	Southbound	-407	0.98	0.98	0.5%
7391	Northbound	-407	0.97	0.98	1.0%
7193	Southbound	-406	0.99	0.99	1.3%
7211	Eastbound	-406	0.99	1.00	5.9%
7193	Northbound	-406	0.99	0.99	2.3%
7211	Westbound	-406	0.99	1.00	7.9%
8150	Westbound	-402	0.98	0.99	4.1%
8150	Eastbound	-402	0.99	1.00	5.5%
7275	Westbound	-355	0.99	0.99	2.0%
7275	Eastbound	-355	0.90	0.95	7.7%
7291	Westbound	-354	0.99	0.99	3.6%
7502	Westbound	-354	0.92	1.00	2.5%
7291	Eastbound	-354	1.00	1.00	5.2%
7502	Eastbound	-354	0.88	1.00	4.1%
7335	Westbound	-335	1.00	1.00	3.1%
7335	Eastbound	-335	1.00	1.00	3.3%
7185	Southbound	-333	0.99	1.00	1.6%
7185	Northbound	-333	0.99	1.00	2.2%
5162	Southbound	-316	0.99	0.99	2.3%
5162	Northbound	-316	1.00	1.00	1.6%
Average Value:			**0.98**	**0.99**	**3.3%**

Visualizing the results aids understanding the anomaly detection methods and how they perform. Figure 4.3 shows how each anomaly detection method identifies and deletes the anomalies in count data. The IQR method is not doing a good job of identifying anomalies in this case since the 75^{th} quantile of hourly counts is greater than the average hourly count in the anomaly group. This is caused by the fact that the anomalous observations are more than 25 percent (i.e. 30.8 percent) of the observed counts (see Table 4.2). K-means method shows a decent performance in identifying anomalies, but it still misses a couple of anomalous observations. TOD & IQR method exhibits very good performance and identifies almost all anomalous observations. This method is applied to remove anomalous data in the subsequent analysis.

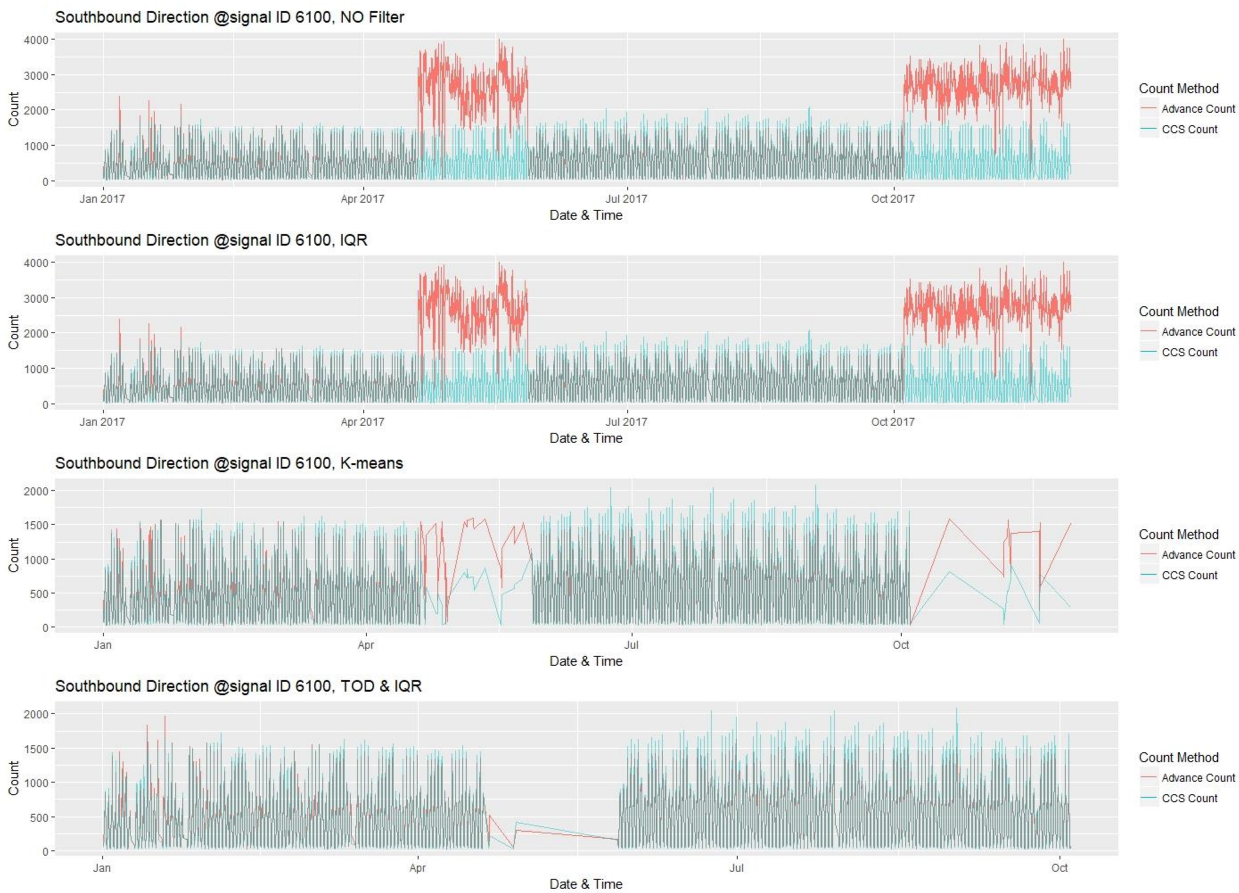

Figure 4.3 Southbound Direction Advance Detector Counts and CCS Counts for Signal ID 6100 for Entire 2017 with Various Filtering Methods (top chart-raw data; second chart-count data following application of the IQR method; third chart-count data following application of the K-means filtering method; fourth chart-count data following application of the IQR + TOD filtering method)

4.4 Accuracy of ATSPM Hourly Counts vs. CCS Hourly Counts

R-squared values are used to determine the accuracy of ATSPM hourly counts. The R-squared value is used due to its simplicity of interpretation and understanding. The accuracy of the ATSPM hourly count was measured under two different settings as follows:

1. ATSPM hourly count without adjustment factor:

$$CCS\ hourly\ count = ATSPM\ hourly\ count + \varepsilon \tag{4-4}$$

2. ATSPM hourly count with adjustment factor:

$$CCS\ hourly\ count = \beta * ATSPM\ hourly\ count + \varepsilon \tag{4-5}$$

where β is the adjustment factor specific for signal and direction. Note that all calculations are performed on the filtered dataset using TOD & IQR anomaly detection method.

Table 4.4 and Table 4.5 show the R-squared with and without adjustment factor and adjustment factor values for Advance and Matrix detectors, respectively.

Table 4.4 R-Squared and Adjustment Value for Advance Detector

Signal ID	Direction	CCS Number	R2 Without Adjustment Factor	R2 With Adjustment Factor	Adjustment Factor
5297	Northbound	-622	0.83	1.00	1.17
5320	Westbound	-620	0.87	0.99	1.13
5011	Southbound	-329	0.80	1.00	1.20
5162	Southbound	-316	0.96	0.99	1.04
7219	Northbound	-332	0.57	0.98	1.43
7185	Southbound	-333	0.70	0.99	1.30
7093	Southbound	-408	0.87	0.95	0.87
7502	Westbound	-354	0.69	1.00	1.31
7275	Westbound	-355	0.59	0.99	0.59
7193	Northbound	-406	0.85	0.99	1.15
7313	Eastbound	-631	0.93	1.00	1.07
7347	Westbound	-335	0.53	0.98	1.47
7524	Southbound	-634	0.98	0.99	1.02
7392	Southbound	-407	0.55	0.98	0.55
7507	Westbound	-633	0.90	0.97	1.10
6100	Southbound	-509	0.95	0.99	1.05
6145	Eastbound	-626	0.98	0.99	0.98
6035	Eastbound	-632	0.72	0.99	1.28
6425	Southbound	-319	0.83	0.89	1.17
6420	Northbound	-350	0.85	0.99	1.15
6256	Westbound	-425	0.79	0.99	0.79
8151	Eastbound	-402	0.96	0.99	0.96
7503	Southbound	-332	0.40	0.96	1.60
7291	Eastbound	-354	-0.25	0.90	2.25
7312	Westbound	-631	0.89	1.00	1.11
7335	Eastbound	-335	0.94	1.00	1.06
7508	Southbound	-633	0.99	0.99	0.99
6185	Westbound	-626	1.00	0.99	1.00
8150	Westbound	-402	0.96	0.99	0.96
7092	Northbound	-408	0.92	1.00	1.08
7211	Eastbound	-406	0.84	0.98	1.16
7391	Northbound	-407	0.94	0.99	1.06
6105	Northbound	-308	0.56	0.86	0.56
Average Value:			**0.79**	**0.98**	**1.11**

Advance detectors without adjustment factors are unreliable since the R-squared values fluctuate significantly across signals and directions. However, Advance detectors with adjustment factors provide a reliable (less than 15% error) estimate of actual hourly counts. On the other hand, Matrix detectors without adjustment factors provide a fairly reliable (less than 20% error) estimate of actual hourly counts.

Table 4.5 R-Squared and Adjustment Value for Matrix Detector

Signal ID	Direction	CCS Number	R2 Without Adjustment Factor	R2 With Adjustment Factor	Adjustment Factor
7313	Westbound	-631	0.88	1.00	1.25
7313	Eastbound	-631	0.96	1.00	1.11
6145	Westbound	-626	0.80	0.98	1.33
6185	Westbound	-626	0.90	0.99	1.21
6145	Eastbound	-626	0.98	0.99	1.02
5320	Westbound	-620	0.95	0.98	1.04
5320	Eastbound	-620	0.95	0.99	1.08
7391	Southbounc	-407	0.91	0.98	1.15
7391	Northbounc	-407	0.93	0.98	1.09
7193	Southbounc	-406	0.90	0.99	0.85
7211	Eastbound	-406	0.91	1.00	0.85
7193	Northbounc	-406	0.98	0.99	1.01
7211	Westbound	-406	0.99	1.00	1.03
8150	Westbound	-402	0.88	0.99	1.20
8150	Eastbound	-402	0.96	1.00	1.10
7275	Westbound	-355	0.92	0.99	1.17
7275	Eastbound	-355	0.77	0.95	1.15
7291	Westbound	-354	0.87	0.99	1.24
7502	Westbound	-354	0.91	1.00	1.21
7291	Eastbound	-354	0.91	1.00	1.18
7502	Eastbound	-354	0.99	1.00	1.01
7335	Westbound	-335	0.99	1.00	1.05
7335	Eastbound	-335	1.00	1.00	1.00
7185	Southbounc	-333	0.98	1.00	1.06
7185	Northbounc	-333	0.96	1.00	1.10
5162	Southbounc	-316	0.92	0.99	1.16
5162	Northbounc	-316	0.97	1.00	1.10
Average Value:			**0.93**	**0.99**	**1.10**

It is worth again pointing out how this research differs from that conducted by Saito, et al, which calculated detector volume adjustment factors for different detector types and lane geometries. In the Saito research, adjustment factors were calculated based on actual hour-long manual counts. The Saito research analyzed data from 18 signals, comparing data for matched hourly periods. The analysis conducted for this UTRAC project is based on an entire year of data (2017) for 11 matched CCS-Matrix locations and 23 matched CCS-Advance locations, and thus is based on a much larger set of sample data over a wider range of conditions.

It is important to note that the R-squared value doesn't consider systematic under/overcounting issue. For example, if a detector is consistently undercounting the actual hourly value by small error, the R-squared value will be high, while the AADT value maybe subject to much larger error due to accumulation of small undercounting errors. This can be seen in Figure 4.4 for signal ID 7313 in the eastbound direction where R-squared values without adjustment factors are high (i.e. 0.96). In these cases, the adjustment factors are further away from the value of 1 (i.e. no adjustment is needed)[1].

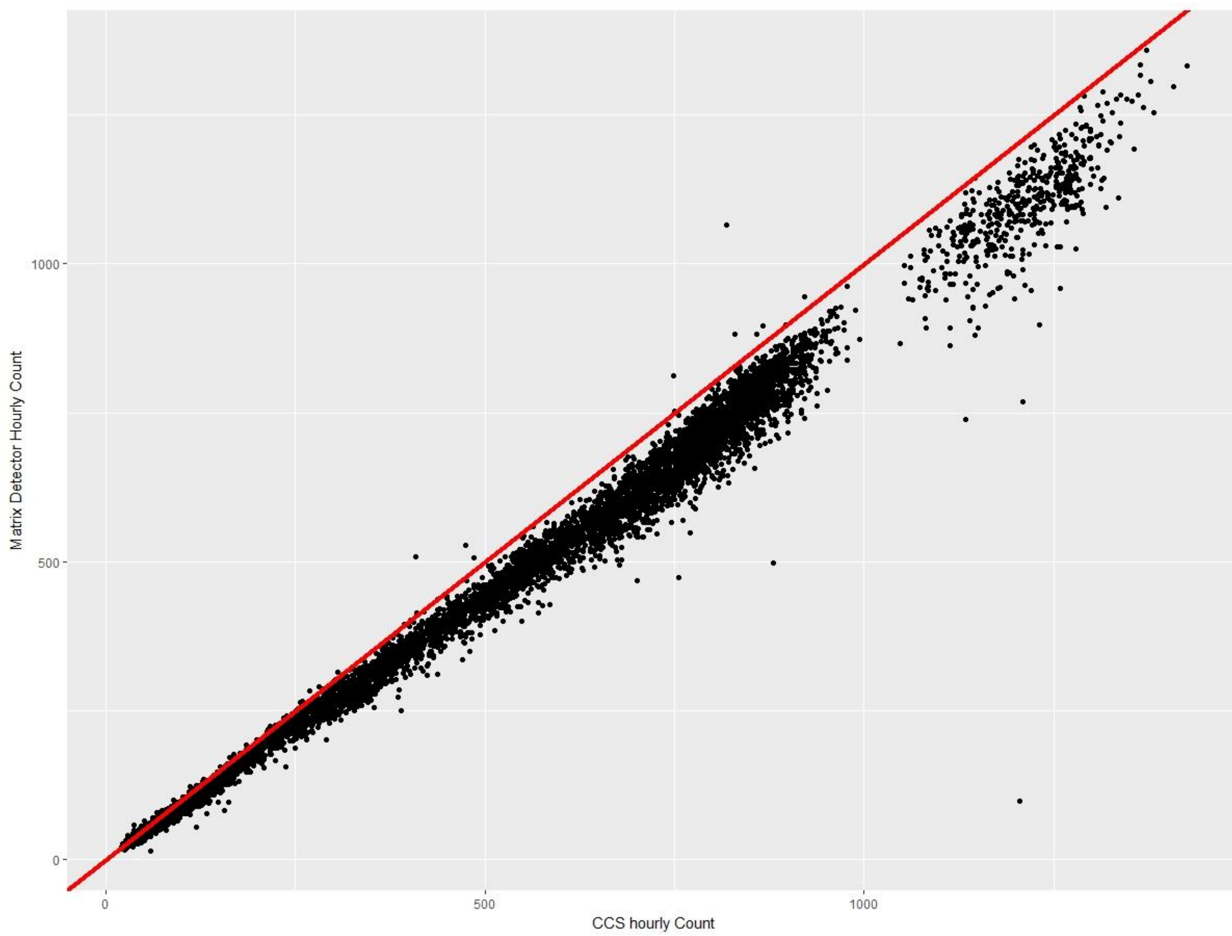

Figure 4.4 Matrix Detector vs. CCS Hourly Counts Scatterplot for Signal ID 7313 in Eastbound Direction

To better understand the accuracy of Matrix detectors without adjustment factors, the Mean Absolute Error (MAE) and Mean Error (ME) were calculated. MAE shows the average value of absolute difference between hourly counts from Matrix detectors and CCS sites. ME

[1] For the purposes of this research the CCS data are held to be ground truth though in reality they may undercount or overcount in some situations.

shows the average value of difference between Matrix detector's and CCS's hourly counts. The detector is systematically under/overcounting if ME and MAE values are close. On the other hand, the detector error is not systematic (i.e. random) if ME is much smaller than MAE and very close to zero. Table 4.6 shows that 67% of Matrix detectors systematically undercount hourly volume, 7% systematically overcount hourly volume, and 26% have non-systematic (i.e. random) error.

Table 4.6 MAE, ME, and Error Type for Matrix Detector Hourly Counts

Signal ID	Direction	CCS Number	MAE	ME	Error Type	Over / Under Counting?
7313	Westbound	-631	114	114	Systematic	Undercounting
7313	Eastbound	-631	56	56	Systematic	Undercounting
6145	Westbound	-626	155	155	Systematic	Undercounting
6185	Westbound	-626	106	104	Systematic	Undercounting
6145	Eastbound	-626	43	16	Random	None
5320	Westbound	-620	26	9	Random	None
5320	Eastbound	-620	26	10	Random	None
7391	Southbounc	-407	63	57	Systematic	Undercounting
7391	Northbounc	-407	54	48	Systematic	Undercounting
7193	Southbounc	-406	105	-99	Systematic	Overcounting
7211	Eastbound	-406	80	-80	Systematic	Overcounting
7193	Northbounc	-406	36	2	Random	None
7211	Westbound	-406	27	16	Random	None
8150	Westbound	-402	99	97	Systematic	Undercounting
8150	Eastbound	-402	57	49	Systematic	Undercounting
7275	Westbound	-355	50	47	Systematic	Undercounting
7275	Eastbound	-355	69	66	Systematic	Undercounting
7291	Westbound	-354	135	128	Systematic	Undercounting
7502	Westbound	-354	119	118	Systematic	Undercounting
7291	Eastbound	-354	101	101	Systematic	Undercounting
7502	Eastbound	-354	31	6	Random	None
7335	Westbound	-335	31	30	Systematic	Undercounting
7335	Eastbound	-335	12	4	Random	None
7185	Southbounc	-333	60	58	Systematic	Undercounting
7185	Northbounc	-333	86	85	Systematic	Undercounting
5162	Southbounc	-316	115	112	Systematic	Undercounting
5162	Northbounc	-316	80	79	Systematic	Undercounting
Average Value:			**72**	**51**	**74% Systematic**	**67% Undercounting, 7% Overcounting, 25% Random**

4.4.1 Impact of Number of Lanes on Detector Accuracy

The distribution of R-squared without adjustment factor and adjustment factor for each lane configuration were plotted to better understand the impact of number of lanes on detector hourly count accuracy. Figure 4.5 and Figure 4.6 show these distributions. The Advance detectors seem to have their best accuracy at two-lane roadways. In addition, the adjustment factor increases as the number of lanes increases. The Advance detectors undercount, neither under or overcount, and overcount for one-lane, two-lane, and three-lane roadways, respectively.

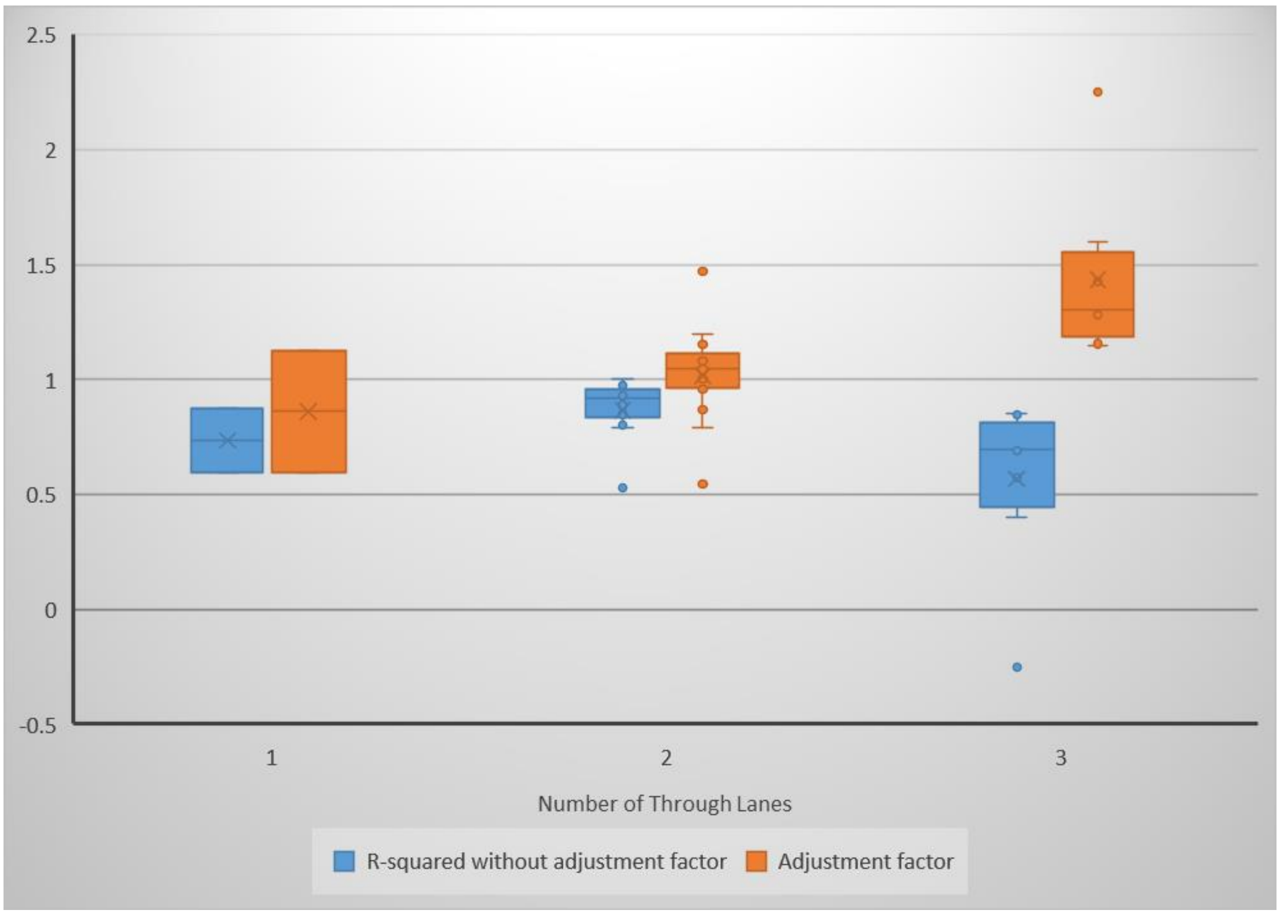

Figure 4.5 R-Squared and Adjustment Factor Distributions for Various Numbers of Through Lanes for Advance Detectors

The accuracy from Matrix detectors is slightly better, and the adjustment factor becomes closer to one as the number of detector channels for the direction increase from three to five.

Note that the comparison for accuracy is comparing one detector type – Matrix – to another detector type – HD.

As mentioned previously, a prior UTRAC study by Saito, et al found that volume counts by the Wavetronix Advance detector tended to undercount actual manual counts, and that the undercounting was more significant as the number of lanes increased. In the current study, comparing counts from the Matrix detector against those collected by the HD detector, the counts increase in accuracy as the number of lanes increases which is likely due to the close correlation of counting errors by both detector types as the number of lanes increases.

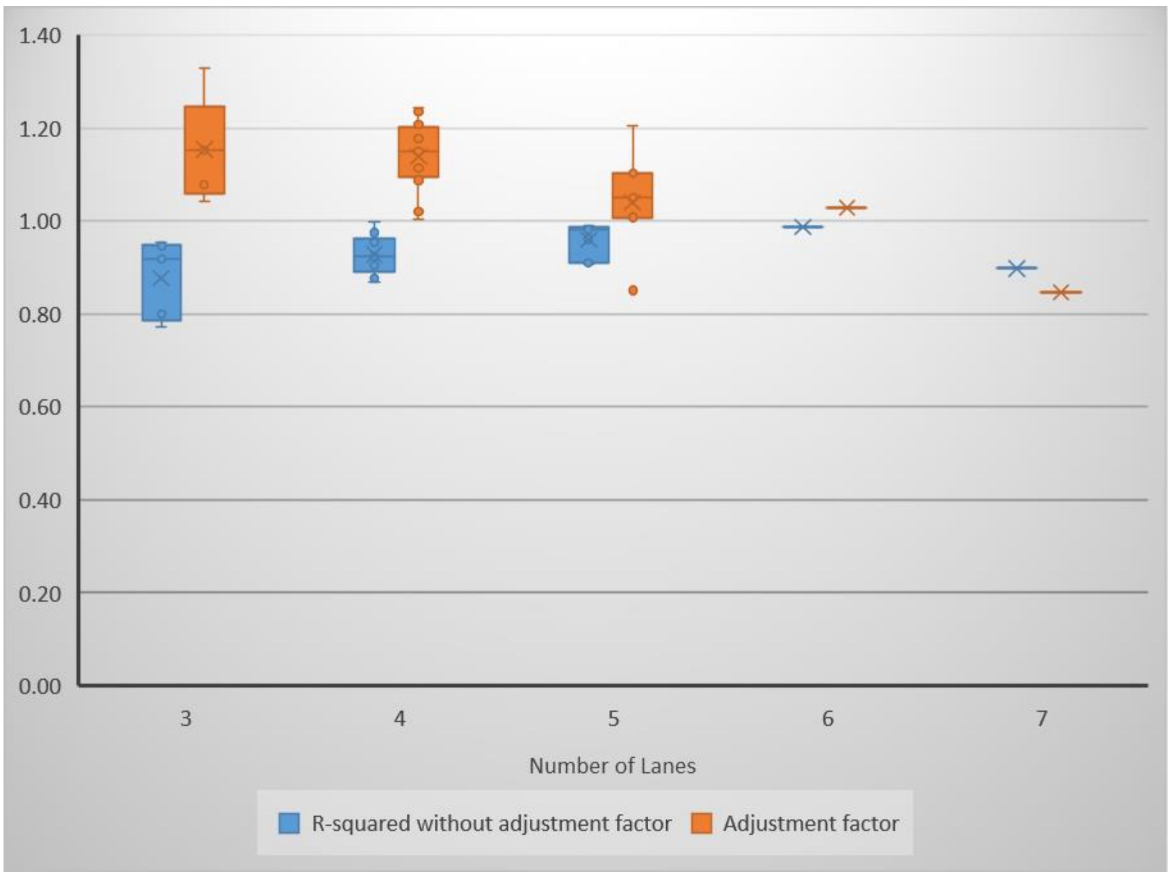

Figure 4.6 R-Squared and Adjustment Factor Distributions for Various Numbers of Detector Channels (lanes) for Matrix Detectors

For each intersection, one directional movement of CCS is measured by one Matrix detector and the other directional movement of a CCS is measured by multiple Matrix detectors. For example, westbound direction for CCS number -620 is measured by signal ID 5320 on westbound direction (i.e. westbound through, westbound right, and westbound left), while the eastbound direction is measured by three different Matrix detectors on eastbound (i.e. eastbound through), northbound (i.e. northbound left), and southbound (i.e. southbound right). Figure 4.7 shows that R-squared (adjustment factor) is generally higher (closer to 1) for single Matrix detector counts compared to multi-Matrix detector counts.

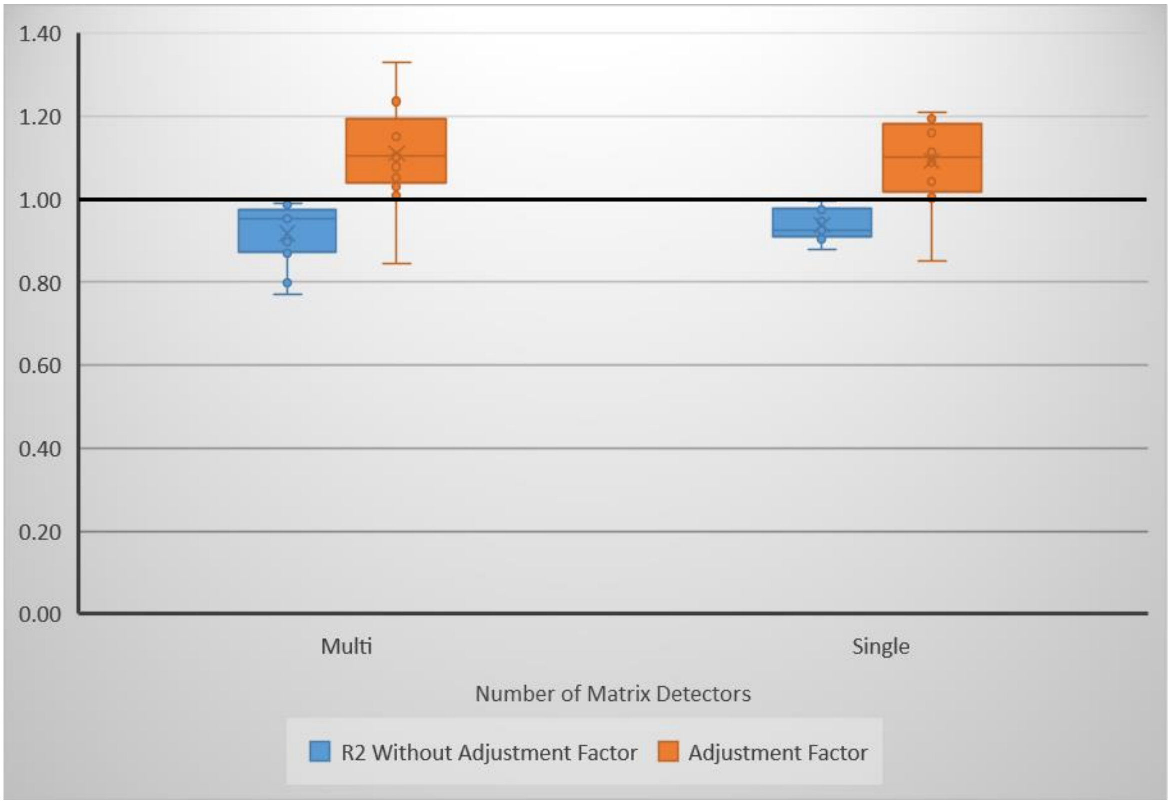

Figure 4.7 R-Squared and Adjustment Factor Distributions for Single or Multiple Matrix Detectors

4.5 Accuracy of Matrix Detectors in Estimating AADT

Volume measurements by Matrix detectors can be used to estimate AADT. AADTs are calculated using the same method as UDOT uses to calculate AADT from CCS count data. The method first calculates the average count for each day of the week in every month (e.g. Monday in January). Then, the average of those seven days in each month is calculated which provides the Monthly Average Daily Traffic (MADT). Finally, the average MADT for 12 months is calculated to create the AADT for a CCS device.

It is important to note that the count data for the entire year may not exist for each signal in the raw Matrix detector data set. This is probably due to the unavailability of the Matrix detector in those periods. For example, if a detector was installed in August, there will be not data for the January to August timeframe. To address this issue, CCS observations for these periods are also removed.

Then the mapped hourly observations that were identified as anomalies in Matrix detector data were removed and imputed by the:

1. Average hourly count for the same hour, day of week, and month,
2. If 1 wasn't available, then average hourly count for the same hour and day of week,
3. If 2 wasn't available, then average hourly count for the same hour.

Table 4.7 shows the percentage of mapped data, CCS AADT, and Matrix detectors' estimated AADT for the 11 signal locations. All prior analysis in this report was based on estimating specific directional volume. The analysis in this section relates to estimating bi-directional volume, or AADT. The Matrix detectors' estimated AADTs are in -21% to +7% (with the average of -9%) range of CCS AADTs. As expected, the Matrix detectors underestimate actual AADT in most cases. The Matrix detectors' estimated AADTs with adjustment factors range in accuracy from a low of -9% to 0% (with the average of -2%) range of actual AADTs.

Table 4.7 Matrix Detector AADT Estimation Accuracy

CCS Number	Signal ID	Mapped Percentage	CCS AADT	Matrix AADT	Matrix/CCS	Matrix/CCS * Adjustment Factor	Note
-631	7313	99%	25,780	21,504	83.4%	98.4%	Thru-Right lane on Westbound
-626	6145	82%	27,895	23,254	83.4%	98.0%	Thru-Right lane on Westbound
-626	6185	82%	14,148	11,691	82.6%	99.8%	Only Westbound direction
-620	5320	93%	10,669	10,031	94.0%	99.7%	
-407	7391	99%	21,614	18,961	87.7%	98.2%	
-406	7193	97%	27,534	29,431	106.9%	99.0%	Thru-Right lane on Southbound
-406	7211	3%	21,685	22,681	104.6%	98.3%	
-402	8150	55%	26,307	22,061	83.9%	96.4%	
-355	7275	98%	16,977	13,319	78.5%	91.0%	Thru-Right lane on Eastbound
-354	7291	26%	32,311	26,211	81.1%	97.9%	Thru-Right lane on Westbound & no detector on Southbound
-354	7502	88%	32,651	29,138	89.2%	98.7%	Thru-Right lane on Eastbound
-335	7335	99%	29,739	28,395	95.5%	98.1%	
-333	7185	97%	43,834	40,025	91.3%	98.7%	Thru-Right lane on Northbound
-316	5162	99%	38,178	33,009	86.5%	97.8%	
Average:		**80%**	**26,380**	**23,551**	**89.2%**	**97.9%**	

As shown in Table 4.7, there are seven cases where approach volumes include through-right movements. This is a source of error that will cause overcounting from Matrix detectors. However, even with this error source, Matrix detectors, across all lanes, are shown to undercount.

4.5.1 Number of Detector Channels and Average Hourly Volume Impact on AADT Estimation Accuracy

The distribution of the ratio of AADTs obtained from Matrix detectors to those estimated from CCSs for both categories of total number of detector channels at intersection (i.e. equal or less than 8 detector channels and more than 8 detectors channels) and two categories of average hourly volume per lane (i.e. equal or less than 100 veh/hr/lane and 101-250 veh/hr/lane) were plotted to better understand their impact on AADT estimation accuracy. Figure 4.8 shows these distributions. AADTs estimated from Matrix detectors are closer to the AADTs estimated from CCS sites (i.e. more accurate) as the total number of detector channels increase or the average volume per hour per lane decreases.

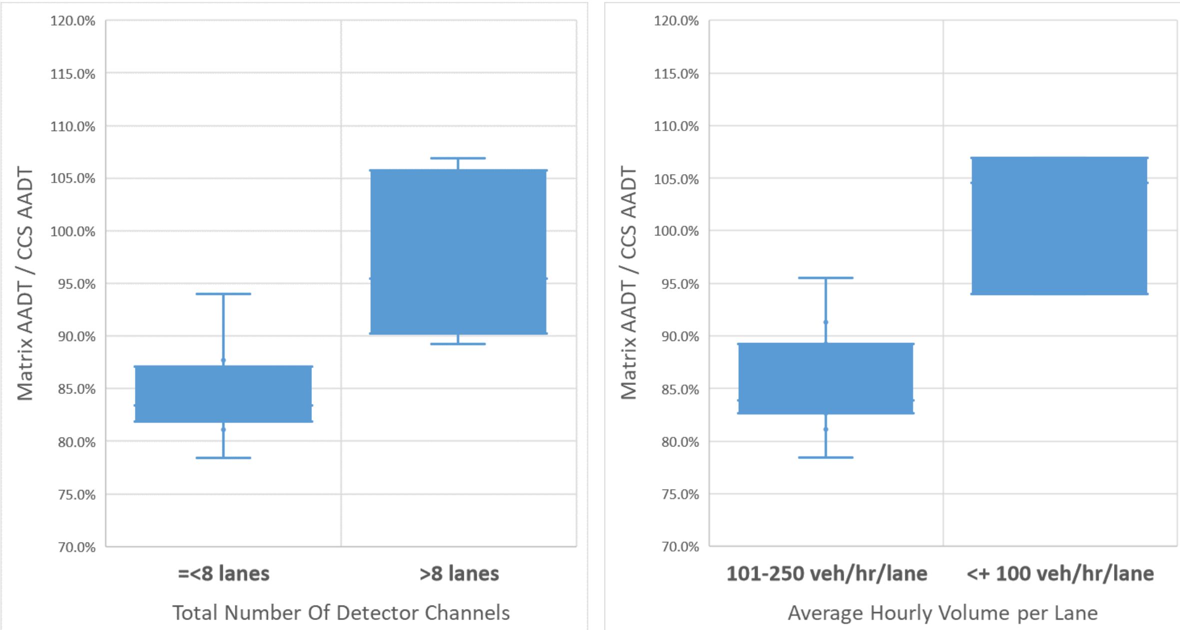

Figure 4.8 Matrix Detector Estimated AADT to CCS Estimated AADT Ratio for Various Number Detector Channels and Average Hourly Volume per Hour per Lane at Intersection

Table 4.8 shows the number of detector channels for the direction with a single detector as well as the direction with multiple detectors. In addition, the total number of detector channels and average volume per hour per lane is provided.

Table 4.8 Lane Configuration for Matrix Intersection

CCS Number	Signal ID	Matrix/ CCS	Single Matrix Detector, Multiple Channels	Multiple Matrix Detector, Multiple Channels	Both Directions Number of Detector Channels	Average Hourly Volume per Lane (veh/hr/lane)
-631	7313	83.4%	4	4	8	134
-626	6145	83.4%	4	3	7	166
-626	6185	82.6%	4		4	147
-620	5320	94.0%	3	3	6	74
-407	7391	87.7%	4	4	8	113
-406	7193	106.9%	5	7	12	96
-406	7211	104.6%	5	6	11	82
-402	8150	83.9%	4	4	8	137
-355	7275	78.5%	3	3	6	118
-354	7291	81.1%	4	4	8	168
-354	7502	89.2%	5	5	10	136
-335	7335	95.5%	4	5	9	138

-333	7185	91.3%	5	5	10	183
-316	5162	86.5%	4	4	8	199
Average:		**89.2%**	**4**	**4**	**8**	**135**

4.6 Accuracy of Matrix Detectors in Estimating Seasonal Factors

Seasonal factors for Matrix detectors and CCS sites' seasonal factors are calculated in the same way, first by averaging the daily traffic for each day (Mondays, Tuesdays, etc.), then averaging all days together to arrive at the MADT for each location. Then the average of seasonal factors was calculated for each functional classification grouping. Figure 4.9 and Figure 4.10 show the scatter plots of Matrix detectors' and CCSs' estimated seasonal factors. The Matrix detectors are very accurate in estimating both monthly and DOW in each month's seasonal factors. The average accuracy of Matrix detectors in estimating the monthly and DOW in each month's seasonal factor for functional classification groupings are 97.56 percent and 96.85 percent, respectively.

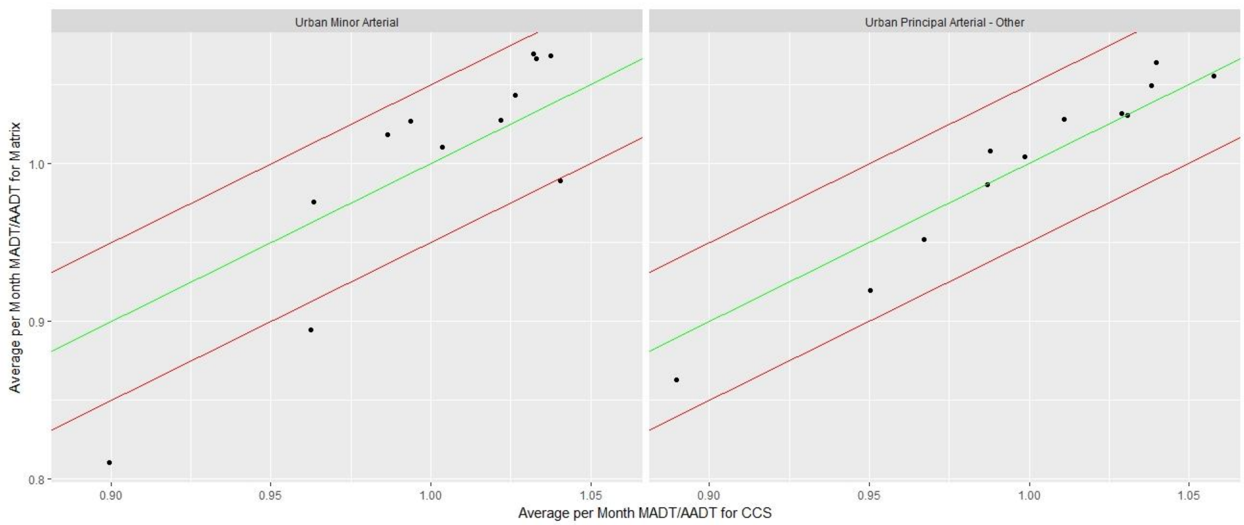

Figure 4.9 Monthly Seasonal Factor Estimation Accuracy for Matrix Detectors (the green line represents 100% accuracy and the red lines represent 5% error bandwidth)

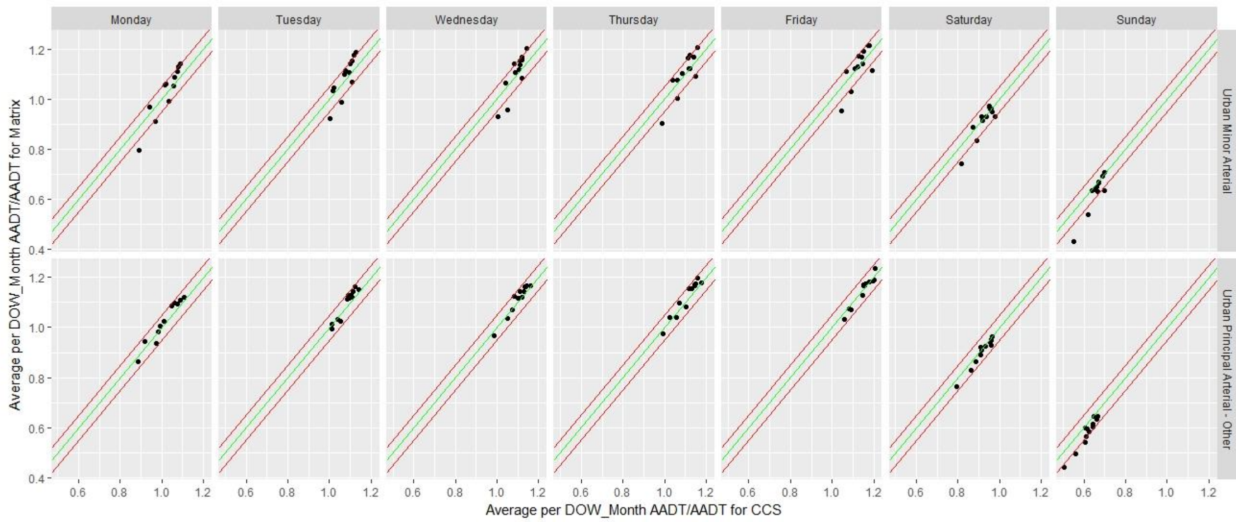

Figure 4.10 DOW in Month Seasonal Factor Estimation Accuracy for Matrix Detectors (the green line represents 100% accuracy and the red lines represent 5% error bandwidth)

4.7 Summary

This chapter has presented the analysis on how traffic volume data measured by Matrix and Advance detectors compares statistically with the traffic volume data obtained from CCS sites, which are themselves based on the Wavetronix sidefire High Definition (HD) detector or on in-pavement inductance loops. The chapter described how the ATSPM sites were mapped to CCS sites, and then described how anomalous volume data from ATSPM sites were identified and filtered. The chapter then described the accuracy of ATSPM volume counts compared to those from CCSs for three types of information:

1. Hourly (Matrix and Advance detectors evaluated)
2. Estimated AADT (Matrix detectors evaluated)
3. Estimated seasonal factors (Matrix detectors evaluated)

For hourly count estimates, Advance detector volume counts reasonably estimated CCS volume counts, with decent statistical fits (r-squared) ranging from 0.70 to 1.00 (average of 0.93). Using a statistically estimated factor adjustment would substantially improve the accuracy of hourly volume estimates from Advance detectors.

Matrix detectors were shown to be more accurate than Advance detectors for hourly count estimates, with an r-square ranging from 0.77 to 1.00 (average of 0.93), for the unadjusted case, and significantly better hourly volume matching using an adjustment factor.

In 2/3 of the cases, Matrix hourly counts systematically undercounted CCS hourly counts which would warrant use of an adjustment factor. In 26% of the cases, there was a random relationship between the Matrix hourly count and the CCS hourly count, indicating that the estimation was sometimes higher and sometimes lower.

Similar findings are reported for the accuracy of Matrix detectors in estimating the AADTs derived from CCS sites, with AADTs estimated from Matrix detectors averaging 89.2% of the AADTs estimated for CCS sites. Regarding estimating seasonal factors, the Matrix detectors are very accurate in estimating both monthly and DOW in each month's seasonal factors. The average accuracy of Matrix detectors in estimating the monthly and DOW in each month's seasonal factor for functional classification groupings are 97.56 percent and 96.85 percent, respectively.

5.0 CONCLUSIONS

5.1 Summary

This research aimed to investigate the possible use of traffic volume data measured from Matrix and Advance detectors from UDOT's ATSPM system to estimate three types of volume-related data currently obtained from UDOT's CCS system. The goal would be to determine whether such volume data obtained from the ATSPM system could be used instead of short-duration traffic counts, thereby saving UDOT significant costs incurred in the short-duration count program. For this purpose, the 11 locations where 27 Matrix detectors exist in proximity of CCSs were identified and 23 locations where 33 Advance detectors exist in proximity to CCS sites, to conduct the comparative analysis. The hourly traffic count data for the entire 2017 period was collected and mapped from the ATSPM system and CCSs for these locations.

Three outlier detection methods were compared for their accuracy to identify anomalous data points in the detectors' data. The accuracy of Matrix and Advance detectors' hourly counts was calculated compared to CCSs' hourly counts as ground-truth data points. Also, adjustment factors for each mapped location for both Matrix and Advance detectors were calculated. The impact of number of lanes and number of detectors on the accuracy and adjustment factors were investigated. Finally, the hourly counts of Matrix detectors and CCSs were used to estimate AADTs and seasonal factors. In addition, the impact of number of detector channels and average volume per hour per lane on AADT estimation accuracy was explored.

5.2 Findings

This section presents research findings from the anomaly identification, hourly count accuracy, AADT estimation accuracy, and seasonal factor estimation accuracy based on the analysis methods employed.

5.2.1 Anomaly Identification

Three anomaly identification methods including IQR, k-means, and TOD & IQR were compared in terms of their accuracy in identifying the erroneous data points. All the methods are solely dependent on ATSPM detectors' data to ensure their applicability to locations where ground-truth data (e.g. CCS) is not available. The TOD & IQR method outperformed the other methods in identifying anomalous data points in hourly data (Table 4.2).

On average 10.5 percent of Advance detectors' hourly counts are anomalous with some detectors having a higher anomaly rate of up to 48.1 percent and some having lower anomaly rates down to 1.6 percent (Table 4.2). This result indicates that anomaly detection must be applied to Advance detectors' hourly counts.

On average 3.3 percent of Matrix detectors' hourly counts are anomalous with some detectors having a slightly higher anomaly rate of up to 7.9 percent and many having a lower anomaly rate down to 0.5 percent (Table 4.3). This result indicates that Matrix detectors' hourly counts are much less prone to malfunctioning compared to Advance detectors' hourly counts. However, we still recommend using an anomaly detection method for Matrix detectors' hourly counts since even small numbers of large anomalies can have a significant impact on the accuracy of Matrix detectors.

5.2.2 Advance and Matrix Detectors' Hourly Count Accuracy

R-squared values are used to determine the accuracy of ATSPM detectors' hourly counts once without an adjustment factor and once with an adjustment factor. The adjustment factor is a singular signal-direction specific value.

The average R-squared values for Advance detectors' hourly counts with and without an adjustment factor are 0.79 (with minimum value of -0.25 and maximum value of 1) and 0.98 (with minimum value of 0.86 and maximum value of 1), respectively (Table 4.4). The large range of R-squared shows that Advance detectors' hourly counts without adjustment factors are not reliable. On the other hand, the Advance detectors' hourly counts become very accurate given the correct adjustment factor. The adjustment factor for each signal and direction can be calculated by comparing Advance detector with short-duration counts.

The average R-squared values for Matrix detectors' hourly counts with and without an adjustment factor are 0.93 (with minimum value of 0.77 and maximum value of 1) and 0.99 (with minimum value of 0.95 and maximum value of 1), respectively (Table 4.5). This result shows that Matrix detectors' hourly counts are much more accurate than Advance detectors' hourly counts. Most Matrix detectors undercount actual traffic which is probably due to occlusion.

5.2.3 Matrix Detectors' AADT Estimation Accuracy

The Matrix detectors' estimated AADTs are in the -21% to +7% (with the average of -9%) range of actual AADTs (Table 4.7 Matrix Detector AADT Estimation Accuracy). As expected, the Matrix detectors underestimate actual AADT in most cases (i.e. 11 out of 14 cases). On average the Matrix detectors' AADT estimation accuracy is 88 percent which is much higher than the current method in place (i.e. FHWA factoring method) which has an accuracy of 78.3 percent (Zhong et al., 2012).

5.2.4 Matrix Detectors' Seasonal Factor Accuracy

The Matrix detectors' estimated seasonal factors accuracy is 97.5 percent and 96.8 percent for the monthly and DOW in each month's seasonal factors for functional classification groupings, respectively. This result indicates that Matrix detectors are a reliable source of seasonal factor estimation and can be used to complement CCS seasonal factor calculations.

5.3 Limitations and Challenges

The only limitation of this study is the limited diversity (only two) of functional classification groupings across test locations. Expanding the study to include short-duration count sites temporarily outfitted with HD detectors for the duration of a comparative study would help address this limitation.

6.0 RECOMMENDATIONS AND IMPLEMENTATION

6.1 Recommendations

This research has compared the accuracy of volume counts obtained from the ATSPM system – from both Matrix and Advance detectors – as compared to volume counts obtained from the UDOT CCS system, which themselves are collected by two other detector types – HD and in-pavement inductance loops.

The impetus for the research is to determine whether the ATSPM volume counts might be used for additional analytics – such as determining seasonal factors – that are otherwise determined from CCS volume counts. Further, there may be cases where the volume data from the ATSPM system is considered accurate enough to be used for short-duration counts. Approximately 1,000 short-duration count sites currently conducted by UDOT are proximate to ATSPM signals and could be studied further to determine whether the accuracy is sufficient to retire the short-duration count site in favor of the ATSPM volume counts obtained from Matrix detectors.

It is recommended that UDOT discuss this potential with FHWA in order to begin the process. FHWA will ultimately need to approve the shift to the ATSPM system to supplant short-duration counts currently conducted with pneumatic road tubes. It is possible that FHWA will want further comparison of the ATSPM volume counts against actual short-duration counts in order to approve of the shift.

6.2 Implementation Plan

The first step in the Implementation Plan is to contact FHWA with the results of this study. Obtaining FHWA input and, ultimately, consent to use ATSPM volume data for either traffic analytics (to calculate seasonal factors) or to supplant short-duration count sites will be necessary to move forward.

Additional studies may be requested and desirable. For example, it would be helpful to identify a small set of short-duration count sites that align well with the ATSPM signal system and which are fully detectorized with Matrix detectors for the relevant traffic flows. Conducting a set of comparisons of actual short-duration counts versus those obtained from the ATSPM system would be a logical next step.

REFERENCES

Saito, M., Chang, D., and Schultz, G. (2015). "Calibration of Automatic Performance Measures – Speed and Volume Data: Volume 1, Evaluation of the Accuracy of Traffic Volume Counts Collected by Microwave Sensors" *UT-15.14*, September 2015.

Saito, M., Sanchez, G. H., and Schultz, G. (2016). "Calibration of Automatic Performance Measures – Speed and Volume Data: Volume 2, Evaluation of the Accuracy of Approach Volume Counts and Speeds Collected by Microwave Sensors" UT-16.05, May 2016.

Zhong, M., Bagheri, E., & Christie, J. (2012). Improving group assignment and AADT estimation accuracy of short-term traffic counts using historical seasonal patterns & Bayesian statistics. Procedia-Social and Behavioral Sciences, 43, 607-617.